Optimal quantization for the pricing of swing options

Olivier Aj Bardou, Sandrine Bouthemy, Gilles Pagès

Optimal quantization for the pricing of swing options

Olivier Bardou* Sandrine Bouthemy† and Gilles Pagès ‡

6th April 2007

Abstract

In this paper, we investigate a numerical algorithm for the pricing of swing options, relying on the so-called optimal quantization method.

The numerical procedure is described in details and numerous simulations are provided to assert its efficiency. In particular, we carry out a comparison with the Longstaff-Schwartz algorithm.

Key words: Swing options, stochastic control, optimal quantization, energy.

Introduction

In increasingly deregulated energy markets, swing options arise as powerful tools for modeling supply contracts [14]. In such an agreement between a buyer and a seller, the buyer always has to pay some amount even if the service or product is not delivered. Therefore, the buyer has to manage his contract by constantly swinging for one state to the other, requiring delivery or not. This is the kind of agreement that usually links an energy producer to a trader. Numerous other examples of energy contracts can be modeled as swing options. From storages [6, 8] to electricity supply [17, 7], this kind of financial device is now widely used. And it has to be noticed that its field of application has recently been extended to the IT domain [12].

Nevertheless, the pricing of swings remains a real challenge. Closely related to a multiple stopping problem [10, 9], swing options require the use of high level numerical schemes. Moreover, the high dimensionality of the underlying price processes and the various constraints to be integrated in the model of contracts based on physical assets such as storages or gas fired power plants increase the difficulty of the problem.

Thus, the most recent technics of mathematical finance have been applied in this context; from trees to Least Squares Monte Carlo based methodology [25, 16, 18], finite elements [26] and duality approximation [20]. But none of these algorithms gives a totally satisfying solution to the valuation and sensitivity analysis of swing contracts.

The aim of this paper is then to introduce and study a recent pricing method that seems very well suited to the question. Optimal Vector Quantization has yet been successfully applied to the valuation of multi-asset American Options [2, 1, 3]. It turns out that this numerical technique is also very efficient in taking into account the physical constraints of swing contracts. For sake of

*Corresponding author. Gaz de France, Research and Development Division, 361 Avenue du Président Wilson - B.P. 33, 93211 Saint-Denis La Plaine cedex. E-mail: olivier-aj.bardou@gazdefrance.com, Phone: +33 1 49 22 54 58, Fax: +33 1 49 22 57 10

†Gaz de France, Research and Development Division, 361 Avenue du Président Wilson - B.P. 33, 93211 Saint-Denis La Plaine cedex. E-mail: sandrine.bouthemy@gazdefrance.com

‡Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail:gpa@ccr.jussieu.fr

simplicity we shall focus on gas supply contracts. After a brief presentation of such agreements and some background on Optimal Quantization methods [22], we show that a careful examination of the properties of the underlying price process can dramatically improve the efficiency of the procedure, as illustrated by several numerical examples.

The paper is organized as follows: in the first section, we describe in details the technical features of the supply contracts (with firm or penalized constraints) with an emphasis on the features of interest in view of a numerical implementation: canonical decomposition and normal form, backward dynamic programming of the resulting stochastic control problem, existence of bang-bang strategies for some appropriate sets of local and global purchased volume constraints. Section 2 is devoted to some background on optimal vector quantization. In Section 3, our algorithm is briefly analyzed and the *a priori* error bound established in the companion paper [4] is stated (as well as the resulting convergence result of the quantized premium toward the true one). In Section 4, numerous simulations are carried out and the quantization method is extensively compared to the well-known least squares regression algorithm "à la Longstaff-Schwartz". An annex explains in details how the price processes we consider in this paper can be quantized in the most efficient way.

1 Introduction to swing options

1.1 Description of the contract

A typical example of swing option is an energy (usually gas or electricity) supply contract with optional clauses on price and volume. The owner of such a contract is allowed to purchase some amount of energy q_{t_k} at time t_k, $k = 0, \ldots, n-1$ until the contract maturity $t_n = T$, usually one year. The purchase price K_k called strike price may be constant or indexed to past values of crude oil. Throughout the paper we will consider that the strike prices are constant and equal to K over the term of the contract. The volume of gas q_{t_k} purchased at time t_i is subject to the local constraint

$$q_{min} \leq q_{t_k} \leq q_{max}.$$

The cumulative volume purchased prior to time $t_k (i.e.$ up to $t_{k-1})$ is defined by $Q_{t_k} = \sum_{\ell=0}^{k-1} q_{t_\ell}$. It must satisfy the following global constraint (at maturity):

$$Q_T = \sum_{k=0}^{n-1} q_{t_k} \in [Q_{min}, Q_{max}].$$

Two approaches can be considered:

– The constraints on the global purchased volumes are firm.

– A penalty is applied if the constraints are not satisfied.

The price at time t of the forward contract delivered at time T is denoted by $F_{t,T}$, $(F_{0,t_k})_{0\leq k\leq n}$ being a deterministic process (the future prices at time 0) available and tradable on the market.

Let $(S_{t_k})_{0\leq k\leq n}$ be the underlying Markov price process defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Note that it can be the observation at time t_k, $k = 0, \ldots, n$ of a continuous time process. Ideally S_t should be the spot price process of the gas $i.e.$ $S_t = F_{t,t}$. However it does not correspond to a tradable instrument which leads to consider in practice the day-ahead contract $F_{t,t+1}$.

We consider its (augmented) natural filtration $\mathcal{F}^S = (\mathcal{F}^S_{t_k})_{0\leq k\leq n}$. The decision sequence $(q_{t_k})_{0\leq k\leq n-1}$ is defined on $(\Omega, \mathcal{A}, \mathbb{P})$ as well and is $\mathcal{F}^S$-adapted, $i.e$ q_{t_k} is $\mathcal{F}_{t_k} = \sigma(S_{t_0}, \ldots, S_{t_k})$ measurable, $k = 0, \ldots, n$. At time t_k the owner of the contract gets $q_{t_k}(S_{t_k} - K)$.

Remark 1.1. The results of this paper can also be applied to every physical asset or contract where the owner reward for a decision q_{t_k} is a function $\psi(t_k, q_{t_k}, S_{t_k})$. In the case of supply contracts, $\psi(t_k, q_{t_k}, S_{t_k}) = q_{t_k}(S_{t_k} - K)$. As for a storage, q_{t_k} represents the amount of gas the owner of the contract decides to inject or withdraw and the profit at each date is then

$$\psi(t_k, q_{t_k}, S_{t_k}) = \begin{cases} -q_{t_k}(S_{t_k} + c_I) & if \quad q_{t_k} \geq 0 \quad (Injection) \\ -q_{t_k}(S_{t_k} - c_W) & if \quad q_{t_k} \leq 0 \quad (Withdrawal) \\ 0 & if \quad q_{t_k} = 0 \quad (Same\ level\ in\ the\ storage) \end{cases}$$

where c_I (resp. c_W) denotes the injection (resp. withdrawal) cost [6].

1.1.1 Case with penalties

We first consider that the penalties are applied at time T if the terminal constraint is violated. For a given consumption strategy $(q_{t_k})_{0 \leq k < n}$, the price is given by at time 0

$$P(0, S_0, 0) = \mathbb{E}\left(\sum_{k=0}^{n-1} e^{-rt_k} q_{t_k}(S_{t_k} - K) + e^{-rT} P_{_T}(S_T, Q_T)|\mathcal{F}_0\right)$$

where r is the interest rate. The function $(x, Q) \mapsto P_{_T}(x, Q)$ is the penalization: $P_{_T}(x, Q) \leq 0$ and $P_{_T}(x, Q)|$ represents the sum that the buyer has to pay if global purchased volume constraints, say $Q_{\min}$ and $Q_{\max}$, are violated. [6] have already investigated this kind of contract.

Then for every non negative $\mathcal{F}_{t_{k-1}}$ measurable random variable Q_{t_k} (representing the cumulated purchased volume up to t_{k-1}), the price of the contract at time $t_k, k = 0, \ldots, n-1$, is given by

$$P(t_k, S_{t_k}, Q_{t_k}) = \operatorname*{ess\,sup}_{(q_{t_\ell})_{k \leq \ell < n}} \mathbb{E}\left(\sum_{\ell=k}^{n-1} e^{-r(t_\ell - t_k)} q_{t_\ell}(S_{t_\ell} - K) + e^{-r(T-t_k)} P_{_T}(S_T, Q_T)|S_{t_k}\right). \quad (1)$$

The standard penalization function is as follows:

$$P_{_T}(x, Q) = -\left(A\, x\, (Q - Q_{min})_- + B\, x\, (Q - Q_{max})_+\right) \quad (2)$$

where A and B are large enough – often equal – positive real constants.

1.1.2 Case with firm constraints

If we consider that constraints cannot be violated, then for every non negative $\mathcal{F}_{t_{k-1}}$ measurable random variable Q_{t_k} defined on $(\Omega, \mathcal{A}, \mathbb{P})$, the price of the contract at time $t_k, k = 0, \ldots, n-1$ is given by:

$$P(t_k, S_{t_k}, Q_{t_k}) = \operatorname*{ess\,sup}_{(q_{t_\ell})_{k \leq \ell \leq n-1} \in \mathcal{A}_{k, Q_{t_k}}^{Q_{min}, Q_{max}}} \mathbb{E}\left(\sum_{\ell=k}^{n-1} e^{-r(t_\ell - t_k)} q_{t_\ell}(S_{t_\ell} - K)|S_{t_k}\right). \quad (3)$$

where

$$\mathcal{A}_{k,Q}^{Q_{min}, Q_{max}} = \left\{(q_{t_\ell})_{k \leq \ell \leq n-1},\ q_{t_\ell} : (\Omega, \mathcal{F}_{t_\ell}, \mathbb{P}) \mapsto [q_{min}, q_{max}],\ \sum_{\ell=k}^{n-1} q_{t_\ell} \in [(Q_{min} - Q)_+, Q_{max} - Q]\right\}.$$

At time 0, we have:

$$P(0, S_0, 0) = \sup_{(q_{t_k})_{0 \leq k \leq n-1} \in \mathcal{A}_{0,0}^{Q_{min}, Q_{max}}} \mathbb{E}\left(\sum_{k=0}^{n-1} e^{-rt_k} q_{t_k}(S_{t_k} - K)\right).$$

Note that this corresponds to the limit case of the contract with penalized constraints when $A = B = +\infty$. Furthermore, one shows that when the penalties $A, B \to +\infty$ in (2), the "penalized" price converge to the "firm" price. This has been confirmed by extensive numerical implementations of both methods. In practice when A, $B \approx 10\,000$ both methods become indistinguishable for usual values of the volume constraints.

1.2 Canonical decomposition and normalized contract

In this section we obtain a decomposition of the payoff of our swing contract (with firm constraints) into two parts, one having a closed form expression. It turns out that this simple decomposition leads to an impressive increase of the precision of the price computation. It plays the role of a variance reducer. Moreover, its straightforward financial interpretation leads to a better understanding of the swing contract.

In fact, we can distinguish a swap part and a normalized swing part:

$$\begin{aligned} P(0,S_0) &= \underbrace{\mathbb{E}\left(\sum_{k=0}^{n-1} q_{min} e^{-rt_k}(S_{t_k}-K)\right)}_{Swap} \\ &\quad + (q_{max}-q_{min}) \underbrace{\sup_{(q_{t_k})\in\mathcal{A}_{[0,1]}^{\tilde{Q}_{min},\tilde{Q}_{max}}(0,0)} \mathbb{E}\left(\sum_{k=0}^{n-1} e^{-rt_k} q_{t_k}(S_{t_k}-K)\right)}_{Normalized\ Contract} \end{aligned} \tag{4}$$

where

$$\mathcal{A}_{[0,1]}^{\tilde{Q}_{min},\tilde{Q}_{max}}(k,Q) = \{(q_{t_\ell})_{k\le\ell\le n-1}, q_{t_\ell} : (\Omega,\mathcal{F}_{t_\ell},\mathbb{P}) \mapsto [0,1], \sum_{\ell=k}^{n-1} q_{t_\ell} \in [\left(\tilde{Q}_{min}-Q\right)_+, \tilde{Q}_{max}-Q]\}$$

and

$$\tilde{Q}_{min} = \frac{(Q_{min}-nq_{min})_+}{q_{max}-q_{min}}, \qquad \tilde{Q}_{max} = \frac{(Q_{max}-nq_{min})_+}{q_{max}-q_{min}}. \tag{5}$$

The price models investigated in the following sections define the spot price as a process centered around the forward curve, and so $\mathbb{E}(S_t) = F_{0,t}$ is known for every $t \in [0,T]$. Thus, the swap part has a closed form given by

$$Swap_0 = q_{min} \sum_{i=0}^{n-1} e^{-rt_k}(F_{0,t_k}-K).$$

The adaptation to contracts with penalized constraints is straightforward and amounts to modifying the penalization function in an appropriate way.

1.3 Dynamic programming equation

In [6], it is shown that, in the penalized problem, optimal consumption is the solution of a dynamic programming equation.

Proposition 1.1. *Assume that for some positive constants p and C, the following inequality holds for any $x > 0$, and $Q \in [n\,q_{min}, n\,q_{max}]$:*

$$|P_{_T}(x,Q)| \le C(1+x^p).$$

Then, there exists an optimal Markovian consumption $q^*(t_k, S_{t_k}, Q_{t_k})$ given by the maximum argument in the following dynamic programming equation:

$$\begin{cases} P(t_k, S_{t_k}, Q_{t_k}) = \max\limits_{q\in[q_{min},q_{max}]} \left\{ q(S_{t_k} - K) + e^{-r(t_{k+1}-t_k)}\mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, Q_{t_k} + q)|S_{t_k}) \right\}, \\ P(T, S_T, Q_T) = P_{_T}(S_T, Q_T). \end{cases} \tag{6}$$

Usually, the function $P_{_T}(x, Q)$ is given by (2). Then, the case with firm constraints corresponds to the limit case where $P_{_T}(x, Q) = (-\infty)\mathbf{1}_{\{x\notin[Q_{\min},Q_{max}]\}}$.

When considering a contract with firm constraints, a more operating form (see [4]) can be the following

$$\begin{aligned} P(t_k, S_{t_k}, Q_{t_k}) = \max\big\{ & q(S_{t_k} - K) + \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, Q_{t_k} + q)|S_{t_k}), \\ & q \in [q_{\min}, q_{\max}],\ Q_{t_k} + q \in [(Q_{\min} - (n-k)q_{\max})_+, (Q_{\max} - (n-k)q_{\min})_+] \big\}. \end{aligned} \tag{7}$$

1.4 Bang Bang consumption

1.4.1 Case with penalties on purchased volumes

[6] showed the following theoretical result.

Theorem 1.2. *Consider the Problem 1 and $P_{_T}(x, Q) = -xP(Q)$, P being a continuously differentiable function. If the following condition holds*

$$\mathbb{P}\left(e^{-rt_k}(S_{t_k} - K) + \mathbb{E}(e^{-rT} S_T P'(Q^*_T)|S_{t_k}, Q^*_{t_k}) = 0\right) = 0,$$

the optimal consumption at time t_k is necessarily of bang-bang type given by

$$\begin{aligned} q^*(t_k, S_{t_k}, Q^*_{t_k}) \quad = \quad & q_{max}\mathbf{1}_{\{e^{-rt_k}(S_{t_k}-K)+\mathbb{E}(e^{-rT}S_T P'(Q^*_T)|S_{t_k},Q^*_{t_k})>0\}} \\ & + q_{min}\mathbf{1}_{\{e^{-rt_k}(S_{t_k}-K)+\mathbb{E}(e^{-rT}S_T P'(Q^*_T)|S_{t_k},Q^*_{t_k})<0\}}. \end{aligned}$$

The above assumption seems difficult to check since it involves the unknown optimal consumption. However, this would be the case provided one shows that the random variable $e^{-rt_k}(S_{t_k} - K) + \mathbb{E}(e^{-rT} S_T P'(Q^*_T)|S_{t_k}, Q^*_{t_k})$ is absolutely continuous as noticed in [6].

1.4.2 Case with firm constraints

In the companion paper [4], we establish some properties of the value function of the swing options viewed *as a function of the global volume constraints* (Q_{min}, Q_{max}). Thanks to (4) one may assume without loss of generality that the contract is normalized, *i.e.* $q_{min} = 0$ and $q_{max} = 1$. We consider the following value function:

$$P(Q_{min}, Q_{max}) = \sup_{(q_{t_k})_{0\le k\le n-1}\in\mathcal{A}_{0,0}^{Q_{min},Q_{max}}} \mathbb{E}\left(\sum_{k=0}^{n-1} e^{-rt_k} q_{t_k}(S_{t_k} - K)\right)$$

defined on the unit (upper) simplex $\{(u, v)\in\mathbb{R}^2,\ 0 \le u \le v \le n\}$.

Proposition 1.3. *The premium function $(Q_{min}, Q_{max}) \mapsto P(Q_{min}, Q_{max})$ is a concave, piecewise affine function of the global purchased volume constraints, affine on elementary triangles $(m, M) + \{(u, v), 0 \le u \le v \le 1\}$, $(m, M) \in \mathbb{N}^2$, $m \le M \le n$ and $(m, M) + \{(u, v), 0 \le v \le u \le 1\}$, $(m, M) \in \mathbb{N}^2$, $m \le M - 1 \le n - 1$ which tile of the unit (upper) simplex.*

Theorem 1.4. *For integral valued global constraints, i.e.* $(Q_{min}, Q_{max}) \in \mathbb{N}^2$*, there always exists a bang-bang optimal strategy i.e. the* a priori $[0,1]$*-valued optimal purchased quantities* $q^*_{t_k}$ *are in fact always equal to* 0 *or* 1*.*

Remark 1.2. This result can be extended in some way to any couple of global constraints when all the payoffs are nonnegative (see [4]). Furthermore, it has nothing to do with the Markov dynamics of the underlying asset and holds in a quite general abstract setting.

An example of the premium function $(Q_{min}, Q_{max}) \mapsto P(Q_{min}, Q_{max})$ is depicted on Figure 1.

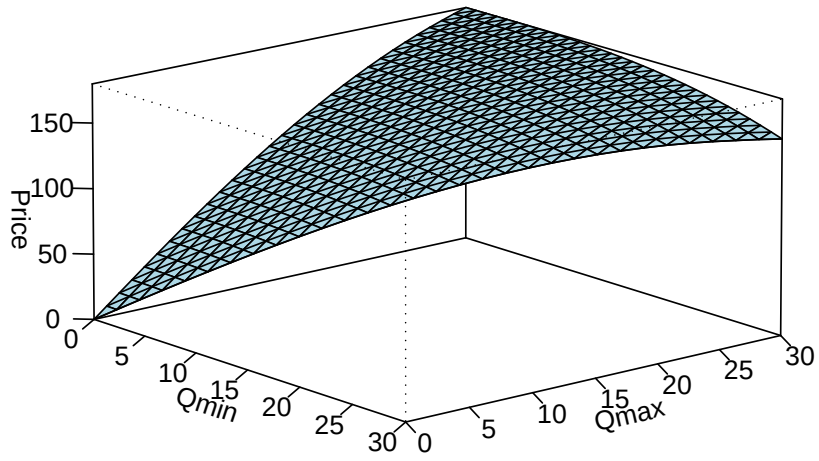

Figure 1: *Value function* $P(Q_{min}, Q_{max})$ *versus the global constraints*

Now we turn to the problem of the numerical evaluation of such contracts. As announced, we focus on an optimal quantization algorithm.

2 Optimal quantization

Optimal Quantization [21, 1, 2, 3] is a method coming from Signal Processing devised to approximate a continuous signal by a discrete one in an optimal way. Originally developed in the 1950's, it was introduced as a quadrature formula for numerical integration in the late 1990's, and for conditional expectation approximations in the early 2000's, in order to price multi-asset American style options.

Let X be an $\mathbb{R}^d$-valued random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Quantization consists in studying the best approximation of X by random vectors taking at most N fixed values $x^1, \ldots, x^N \in \mathbb{R}^d$.

Definition 2.1. Let $x = (x^1, \ldots, x^N) \in (\mathbb{R}^d)^N$. A partition $(C_i(x))_{i=1,\ldots,N}$ of $\mathbb{R}^d$ is a *Voronoi tessellation* of the N-quantizer x (or *codebook*; the term *grid* being used for $\{x^1, \ldots, x^N\}$) if, for every $i \in \{1, \ldots, N\}$, $C_i(x)$ is a Borel set satisfying

$$C_i(x) \subset \{\xi \in \mathbb{R}^d, |\xi - x^i| \leq \min_{i \neq j} |\xi - x^j|\}$$

where $|\,.\,|$ denotes the canonical Euclidean norm on $\mathbb{R}^d$.

The nearest neighbour projection on x induced by a Voronoi partition is defined by

$$\mathrm{Proj}_x : y \in \mathbb{R}^d \mapsto x^i \;\text{ if }\; y \in C_i(x).$$

Then, we define an *x-quantization* of X by

$$\hat{X}^x = \mathrm{Proj}_x(X).$$

The *pointwise error* induced when replacing X by $\widehat{X}^x$ is given by $|X - \widehat{X}^x| = d(X, \{x^1, \ldots, x^N\}) = \min_{1 \leq i \leq N} |X - x^i|$. When X has an absolutely continuous distribution, any two x-quantizations are $\mathbb{P}$-*a.s.* equal.

The quadratic *mean quantization error* induced by the the N-tuple $x \in \mathbb{R}^d$ is defined as the quadratic norm of the pointwise error *i.e.* $\|X - \hat{X}^x\|_2$.

We briefly recall some classical facts about theoretical and numerical aspects of Optimal Quantization. For details we refer *e.g.* to [15, 22].

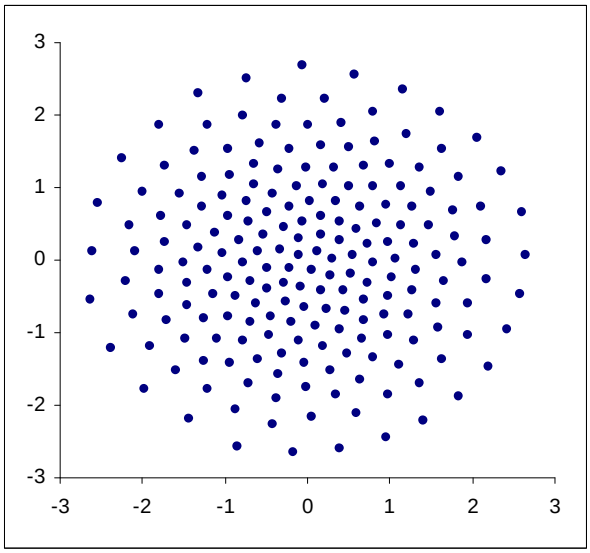

Figure 2: *Optimal quadratic quantization of the normal distribution $\mathcal{N}(0, I_2)$, $N = 200$.*

Theorem 2.1. *[15] Let $X \in L^2(\mathbb{R}^d, \mathbb{P})$. The quadratic quantization error function*

$$x = (x^1, \ldots, x^N) \longmapsto \mathbb{E}(\min_{1 \leq i \leq N} |X - x^i|^2) = \|X - \hat{X}^x\|_2$$

reaches a minimum at some quantizer x^*. Furthermore, if the distribution $\mathbb{P}_X$ has an infinite support then $x^{*,(N)} = (x^{*,1}, \ldots, x^{*,N})$ has pairwise distinct components and $N \mapsto \min_{x \in (\mathbb{R}^d)^N} \|X - \hat{X}^x\|_2$ is decreasing to 0 as $N \uparrow +\infty$.

Figure 2 shows a quadratic optimal quantization grid for a bivariate normal distribution $\mathcal{N}(0, I_2)$. The convergence rate to 0 of optimal quantization error is ruled by the so-called Zador Theorem.

Theorem 2.2. *[15] Let $X \in L^{2+\delta}(\mathbb{P}), \delta > 0$, with $\mathbb{P}_X(d\xi) = \varphi(\xi)\lambda_d(d\xi) + \nu(d\xi)$, $\nu \perp \lambda_d$ (λ_d Lebesgue measure on $\mathbb{R}^d$). Then*

$$\lim_{N\to+\infty} (N^{\frac{2}{d}} \min_{x\in(\mathbb{R}^d)^N} \|X - \hat{X}^x\|_2) = J_{2,d} \left(\int_{\mathbb{R}^d} \varphi^{\frac{d}{d+2}} d\lambda_d \right)^{1+\frac{2}{d}}.$$

The true value of $J_{2,d}$ is unknown as soon as $d \geq 3$. One only knows that $J_{2,d} = \frac{d}{2\pi e} + o(d)$.

Zador's Theorem implies that $\|X - \hat{X}^{x^{*,(N)}}\|_2 = O(N^{-\frac{1}{d}})$ as $N \to +\infty$.

Proposition 2.3. *[21, 22] Any L^2-optimal quantizer $x \in \mathbb{R}^d$ satisfy the following stationarity property*

$$\mathbb{E}(X|\hat{X}^x) = \hat{X}^x.$$

In particular, for any stationary quantizer $\mathbb{E}(X) = \mathbb{E}(\hat{X}^x)$.

The random vector $\hat{X}^x$ takes its value in a finite space $\{x^1, \ldots, x^N\}$, so for every continuous functional $f : \mathbb{R}^d \to \mathbb{R}$ with $f(X) \in L^2(\mathbb{P})$, we have

$$\mathbb{E}(f(\hat{X}^x)) = \sum_{i=1}^{N} f(x^i)\mathbb{P}(X \in C_i(x))$$

which is the quantization based quadrature formula to approximate $\mathbb{E}(f(X))$ [21, 22]. As $\hat{X}^x$ is close to X , it is natural to estimate $\mathbb{E}(f(X))$ by $\mathbb{E}(f(\hat{X}^x))$ when f is continuous. Furthermore, when f is smooth enough, on can upper bound the resulting error using $\|X - \hat{X}^x\|_2$, or even $\|X - \hat{X}^x\|_2^2$ (when the quantizer x is stationary).

The same idea can be used to approximate the conditional expectation $\mathbb{E}(f(X)|Y)$ by $\mathbb{E}(f(\hat{X})|\hat{Y})$, but one also needs the transition probabilities:

$$\mathbb{P}(X \in C_j(x)|Y \in C_i(y)).$$

The application of this technique to the quantization of spot price processes is discussed in details in the Annex, page 22.

3 Pricing swing contracts with optimal quantization

3.1 Description of the algorithm (general setting)

In this section we assume that $(S_{t_k})_{0\leq k\leq n}$ is a Markov process. For sake of simplicity, we consider that there is no interest rate. We also consider a normalized contract, as defined in Section 1.2.

In the penalized problem, the price of the swing option is given by the following dynamic programming equation (see Equation 6):

$$\begin{cases} P(t_k, S_{t_k}, Q_{t_k}) = \max\limits_{q\in\{0,1\}} [q(S_{t_k} - K) + \mathbb{E}(P(t_{k+1}, S_{t_{k+1}}, Q_{t_k} + q)|S_{t_k})] \\ \\ P(T, S_T, Q_T) = P_T(S_T, Q_T) \end{cases}$$

where $t_k = k\Delta, k = 0, \ldots, n$, $\Delta = \frac{T}{n}$.

The bang-bang feature of the optimal consumption (see Section 1.4) allows us to limit the possible values of q in the dynamic programming equation to $q \in \{0,1\}$. At time t_k, possible values of the cumulative consumption are

$$Q_{t_k}^{\ell} = \ell, 0 \leq \ell \leq k. \tag{8}$$

At every time t_k we consider a(n optimized) N_k-quantization $\widehat{S}_{t_k} = \widehat{S}_{t_k}^{x_k^{(N)}}$, $k = 0, \ldots, n$ based on an optimized quantization N_k-tuple (or grid) $x_k^{(N_k)} := \left(s_k^1, \ldots, s_k^{N_k}\right)$ of the spot S_{t_k}.

The modeling of the future price by multi-factor Gaussian processes with memory (see Section 4 for a toy example) implies that $(S_t)_{t\in[0,T]}$ is itself a Gaussian process. Then the quantization of S_{t_k} can be obtained by a simple dilatation-contraction (by a factor $\mathrm{StD}(S_{t_k})$) from optimal quantization grids of the (possibly multivariate) normal distribution, to be downloaded on the website [23]

`www.quantize.maths-fi.com`

Then we compute the price at each time t_k, for all points on the corresponding grid, and for all the possible cumulative consumptions:

$$\begin{cases} P(t_k, s_k^i, \hat{Q}_{t_k}) = \max\limits_{q\in\{0,1\}} \left[q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, \hat{Q}_{t_k} + q)|\hat{S}_{t_k} = s_k^i)\right] \\ \qquad\qquad i = 1, \ldots, N_k, \\ P(T, s_T^i, \hat{Q}_T) = P_T(s_T^i, \hat{Q}_T),\ i = 1, \ldots, N_n. \end{cases} \tag{9}$$

When considering a contract with firm constraints, we need to compute the price at each time t_k, for all the points of the quantization grid of the spot price, and for all the admissible cumulative consumptions (See Figure 3)

$$Q_{t_k}^{\ell} = \ell + (Q_{\min} - (n-k+1))_+, \ell = 0, ..., \min(k, Q_{\max}) - (Q_{\min} - (n-k+1))_+ . \tag{10}$$

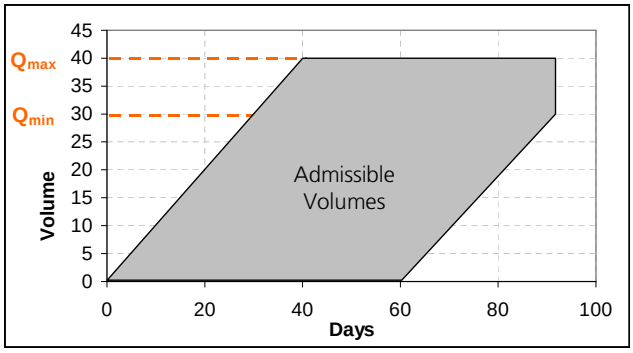

Figure 3: *Volume Constraints*

Using the bang-bang feature (See section 1.4) and the dynamic programming principle (Equation 7), this price is given by

$$P(t_k, s_k^i, \hat{Q}_{t_k}) = \max\Big\{q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, \hat{Q}_{t_k} + q)|\hat{S}_{t_k} = s_k^i), \\ q \in \{0,1\},\ \hat{Q}_{t_k} + q \in [(Q_{\min} - (n-k))_+, Q_{\max}]\Big\}. \tag{11}$$

Since $\hat{S}_{t_k}$ takes its values in a finite space, we can rewrite the conditional expectation as:

$$\mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, Q)|\hat{S}_{t_k} = s_k^i) = \sum_{j=1}^{N_{k+1}} P(t_{k+1}, s_{k+1}^j, Q)\pi_k^{ij}$$

where

$$\pi_k^{ij} = \mathbb{P}(\hat{S}_{t_{k+1}} = s_{k+1}^j | \hat{S}_{t_k} = s_k^i)$$

is the *quantized transition probability* between times t_k and t_{k+1}. The whole set of quantization grids equipped with the transition matrices make up the so-called "quantization tree". The transition weights (π_k^{ij}) matrices are the second quantity needed to process the quantized dynamic programming principle (9 or 11). A specific fast parallel quantization procedure has been developed in our multi-factor Gaussian framework to speed up (and parallelize) the computation of these weights (see Annex). In a more general framework, one can follow the usual Monte Carlo approach described in [1] to compute the grids and the transition of the global quantization tree.

When (S_{t_k}) has no longer a Markov dynamics but appears as a function of a Markov chain (X_k): $S_{t_k} = f(X_k)$, one can proceed as above, except that one has to quantize (X_k). Then the dimension of the problem (in term of quantization) is that of the structure process (X_k).

Of course one can always implement the above procedure formally regardless of the intrinsic dynamics of (S_{t_k}). This yields to a drastic dimension reduction (from that of X downto that of (S_{t_k})). Doing so, we cannot apply the convergence theorem (see Section 3.3) which says that in a Markovian framework the premium resulting from (9) or (11) will converge toward the true one as the size of the quantization grid goes to infinity.

This introduces a methodological residual error that can be compared to that appearing in [5] algorithm for American option pricing. However, one checks on simulations that this residual error turns out to be often negligible (see Section 4.3).

3.2 Complexity

The first part of the algorithm consists in designing the quantization tree and the corresponding weights. The complexity of this step is directly connected to the size of the quantization grids chosen for the transitions computation in 1-dimension, or to the number of Monte Carlo simulations otherwise. However those probabilities have to be calculated once for a given price model, and then several contracts can be priced on the same quantization tree. So we will mainly focus on the complexity of the pricing part.

We consider a penalized normalized contract, *i.e* $q_{min} = 0$ and $q_{max} = 1$. The implementation of the dynamic programming principle requires three interlinked loops. For each time step k (going backward from n to 0), one needs to compute for all the points s_k^i, $i = 1, \dots, N_k$ of the grid and for every possible cumulative consumption $Q_{t_k}^\ell$ $(0 \le \ell \le k)$ (see (8)) the functional

$$\max_{q \in \{0,1\}} [q(s_k^i - K) + \mathbb{E}(P(t_{k+1}, \hat{S}_{t_{k+1}}, Q_{t_k} + q) | \hat{S}_{t_k} = s_k^i)]$$

which means computing twice a sum of N_{k+1} terms.

Hence, the complexity is proportional to

$$\sum_{k=0}^{n-1} (k+1) N_k N_{k+1}.$$

In the case where all layers in the quantization tree have the same size, *i.e* $N = N_k, \forall k = 1, \dots, n$, the complexity is proportional to $\frac{n^2 N^2}{2}$. This is not an optimal design but *only one grid needs to be stored*. It is possible to reduce the algorithm complexity by optimizing the grid sizes N_i [1](with the constraint $\sum_k N_k = nN$), but it costs more memory space.

[1] To minimize the complexity, set $N_k \approx \frac{2nN}{(k+1)log(n)}$, $k = 0, \dots, n-1$, which leads to a global complexity proportional to $\frac{4n^2N^2}{\log(n)}$

In the case of firm constraints, the dynamic programming principle (11) has to be computed for every admissible cumulative consumption, *i.e* for every $Q_{t_k}^{\ell}$ ($0 \leq \ell \leq \min(k-1, Q_{\max}) - (Q_{\min} - (n-k+1))_+$, see (10)). The complexity is proportional to

$$\sum_{k=0}^{n-1} (\min(k, Q_{\max}) - (Q_{\min} - (n-k+1))_+ + 1) N_k N_{k+1}.$$

The complexity in the case of firm constraints is lower than the one for a penalized problem, and depends on the global constraints $(Q_{\min}, Q_{\max})$. But the implementation is easier in the case of a penalized problem, because one does not need to check if the cumulative consumption volume is admissible. Both approaches have been numerically tested and results are indistinguishable for large enough penalties. For the implementation readiness, the approach with penalties has been adopted.

In order to reduce the complexity of the algorithm, one usually prunes the quantization tree. In most examples, at each layer k, many terms of the transition matrix $(\pi_{ij}^k)_{i,j}$ are equal to 0 or negligible. So while the transition probabilities are estimated, all the transitions that are not visited are deleted. This step is important because it allows to reduce significantly the algorithm complexity.

In practice we can even neglect transitions whose probability is very low, say less than 10^{-5}.

3.3 Convergence

In [4] is proved an error bound for the pricing of swing options by optimal quantization.

Let $P_0^n(Q)$ denote the price of the swing contract at time 0. n is the number of time step, and $Q = (Q_{min}, Q_{max})$ is the global constraint. We consider a contract with normalized local constraints, *i.e* $q_{min} = 0$ and $q_{max} = 1$. The "quantized" price $\hat{P}_0^n(Q)$ is the approximation of the price obtained using optimal quantization.

Proposition 3.1. *Assume there is a real exponent $p \in [1, +\infty)$ such that the (d-dimensional) Markov structure process $(X_k)_{0 \leq k \leq n-1}$ satisfies*

$$\max_{0 \leq k \leq n-1} |X_k| \in L^{p+\eta}(\mathbb{P}), \eta > 0.$$

At each time $k \in \{0, \ldots, n-1\}$, we implement a (quadratic) optimal quantization grid x^N of size N of X_k. Then

$$\| \sup_{Q \in T^+(n)} |P_0^n(Q) - \hat{P}_0^n(Q)| \|_p \leq C \frac{n}{N^{\frac{1}{d}}}$$

where $T^+(n) := \{(u, v), 0 \leq u \leq v \leq n\}$ is the set of admissible global constraints (at time 0).

In fact this error bound turns out to be conservative and several numerical experiments, as those presented in Section 4, suggest that in fact the true rate (for a fixed number n of purchase instants) behaves like $O(N^{-\frac{2}{d}})$.

4 Numerical experiments

In this section the same grid size has been used at each time step, *i.e.* we always have $N_k = N, k = 0, \ldots, n$. The results have been obtained by implementing the penalized problem and using the canonical decomposition (see Section 1.2).

4.1 The one factor model

Swing options are often priced using the least squares regression method "à la Longstaff-Schwartz" [19]. This section aims to compare our numerical results to those obtained with Longstaff-Schwartz method. We consider a one factor model, which corresponds to a one dimensional Markov structure process.

4.1.1 Quantization tree for a one dimensional structure process

We consider the following diffusion model for the forward contracts $(F_{t,T})_{0\leq t\leq T}$:

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma e^{-\alpha(T-t)} dW_t$$

where W is a standard Brownian motion. It yields:

$$S_t = F_{0,t} \exp\left(\sigma \int_0^t e^{-\alpha(t-s)} dW_s - \frac{1}{2}\Lambda_t^2\right)$$

where

$$\Lambda_t^2 = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}).$$

Denote $X_k = \int_0^{k\Delta} e^{-\alpha(t-s)} dW_s$. The structure process $(X_k)_{k\geq 0}$ can be quantized using the fast parallel quantization method described in the Annex (page 22). Let $x_k^{(N)}$ denote an (optimal) quantization grid of X_k of size N. We have to compute for every $k \in \{0, \ldots, n-1\}$, and every $(i,j) \in \{1, \ldots, N\}^2$, the following (quantized transition) probabilities:

$$p_k^{ij} = \mathbb{P}(\eta_1 \in C_i(x_k^{(N)}); \alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(x_{k+1}^{(N)})) \tag{12}$$

where $(\eta_1, \eta_2) \sim \mathcal{N}(0, I_2)$, and α_k and β_k are scalar coefficients that can be explicited.

This can be done by using quantization again and importance sampling as presented in the Annex, page 22 (see (21)).

4.1.2 Comparison with the regression method

We first use the following parameters for the one factor model:

$$\sigma = 70\%,\ \alpha = 4,\ F_{0,t_k} = 20,\ k = 0, \ldots, n.$$

The following tables present the results obtained with Longstaff Schwartz and optimal quantization, for different strike values. 1000 Monte-Carlo sample paths have been used for Longstaff-Schwartz method and the confidence interval of the Monte Carlo estimate is given in the table. A 100-point grid has been used to quantize the spot price process, and the transitions have been computed with a 500-point grid. The local volume constraints q_{min} and q_{max} are set to 0 and 6 respectively .

We first consider a case without constraints (Table 1), which means that the swing option is a strip of calls, whose price is easily computed with the Black-Scholes formula.

Table 2 presents the results obtained with the global constraints $Q_{min} = 1300$ and $Q_{max} = 1900$. Volume constraints are presented on Figure 4.

	$K = 5$	$K = 10$	$K = 15$	$K = 20$
Longstaff-Schwartz	[32424,33154]	[21360,22127]	[11110,11824]	[3653,4109]
10 point grid	32726	21806	11311	3905
20 point grid	32751	21834	11367	3943
50 point grid	32759	21843	11380	3964
100 point grid	32759	21843	11380	3964
200 point grid	32761	21845	11382	3967
Theoretical price	32760	21844	11381	3966

Table 1: *Comparison for a call strip (no global constraints)*

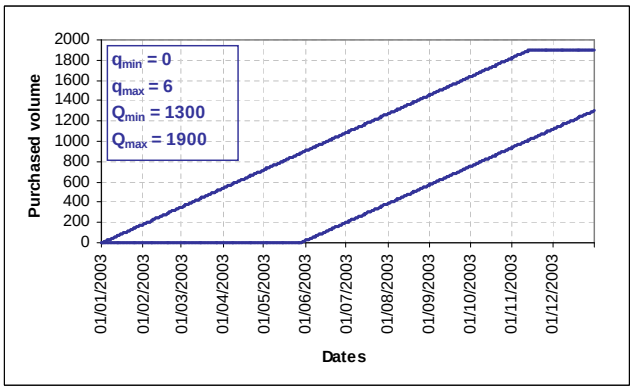

Figure 4: *Consumption constraints*

	$K = 5$	$K = 10$	$K = 15$	$K = 20$
Longstaff-Schwartz	[29068,;29758]	[19318,;19993]	[10265,;10892]	[2482,;3038]
10 point grid	29696	20216	10981	3067
20 point grid	29494	20018	10841	2863
50 point grid	29372	19895	10729	2718
100 point grid	29348	19872	10704	2687
200 point grid	29342	19866	10698	2680

Table 2: *Comparison with constraints*

The results seem consistent for both methods, the price given by quantization always belongs to the confidence interval of the Longstaff-Schwartz method. One can note that it is true even for small grids, which means that quantization gives quickly a good price approximation. Moreover, the price given by quantization is very close of the theoretical price in the case of a call strip.

4.1.3 Execution time

In this section are compared the execution times to price swing options using optimal quantization and Longstaff-Schwartz method.

The size of the quantization grid is 100 for the pricing part and 200 for the transitions computation. And 1000 Monte Carlo simulations are used. The maturity of the contract is one year.

The computer that has been used has the following characteristics:

Processor: Celeron; CPU 2,4 Ghz; 1,5 Go of RAM; Microsoft Windows 2000.

The execution times given in Table 3 concern the pricing of one contract, which yields the building of the quantization tree and the pricing using dynamic programming for quantization.

Longstaff-Schwartz	Quantization: Quantization tree building + Pricing	Quantization: Pricing only
160 s	65 s	5 s

Table 3: *Execution time for the pricing of one contract*

If we consider the pricing of several contracts, there is no need for re computing the quantization tree if the underlying price model has not changed. That is why quantization is really faster than Longstaff-Schwartz in this case, as one can note from the results presented in Table 4.

Longstaff-Schwartz	Quantization
1600 s	110 s

Table 4: *Execution time for the pricing of 10 contracts*

4.1.4 Sensitivity Analysis

When contracts such as swing options ought to be signed, negotiations usually concern the volume constraints. That is why the valuation technique has to be very sensitive and coherent to constraints variation. In this section we will compare the sensibility to global constraints for Longstaff-Schwartz method and optimal quantization.

Figure 5 represents the price of the contract with regards to the global constraints Q_{min} and Q_{max}, and Figure 6 represents the price versus Q_{max} for a fixed value of Q_{min} equal to 1300.

One can notice that the surface obtained with optimal quantization is very smooth. If Q_{max} increases, the price increases. However, it is not always true with Longstaff-Schwartz because of the randomness of the method, and the limited number of Monte Carlo simulations imposed by the dimension of the problem and the number of time steps.

New Monte Carlo simulations are done for each different contract, *i.e* each time Q_{min} or Q_{max} varies. Of course, the same simulations could be used to price all the contracts, but unfortunately these simulations could be concentrated in the distribution queues and give a price far from the

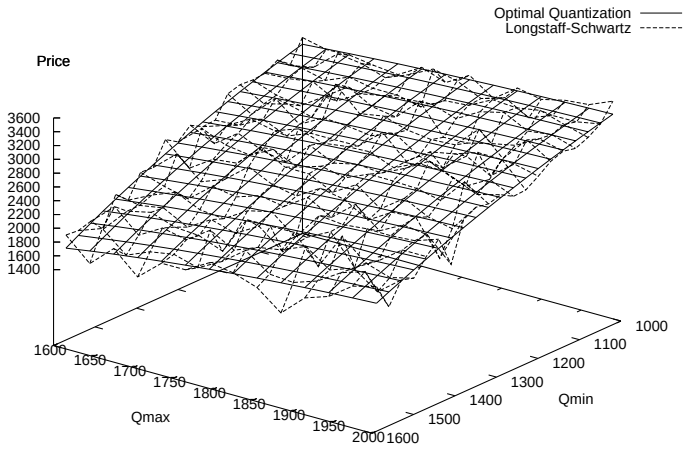

Figure 5: *Sensitivity to global constraints*

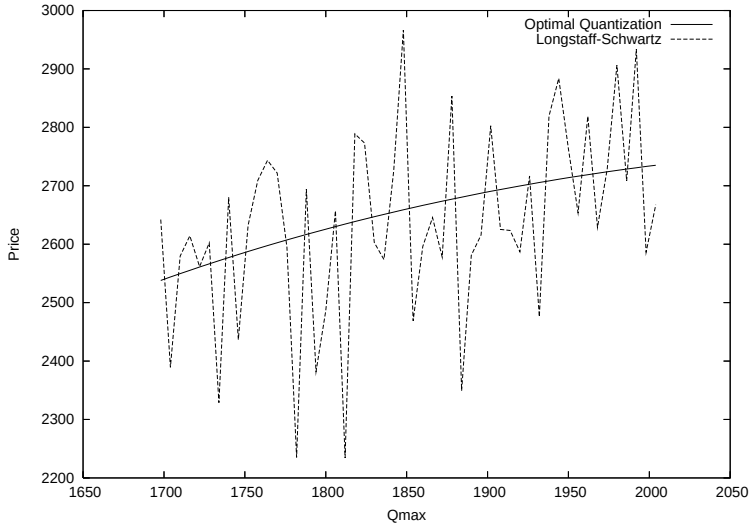

Figure 6: *Sensibility to* Q_{max}

real one for all the contracts. As concerns quantization, the grid is build in order to give a good representation of the considered random variable. One of the great advantages of optimal quantization over Monte-Carlo is that this first algorithm always approximates the whole distribution of the payoff meanwhile it can take a while before Monte-Carlo explores some parts of it.

4.1.5 Convergence

In this section we will study the convergence of the quantization method. We focus on the convergence of the pricing part of the algorithm.

We consider a one year maturity contract with the volume constraints depicted on Figure 4, and the daily forward curve depicted on Figure 7.

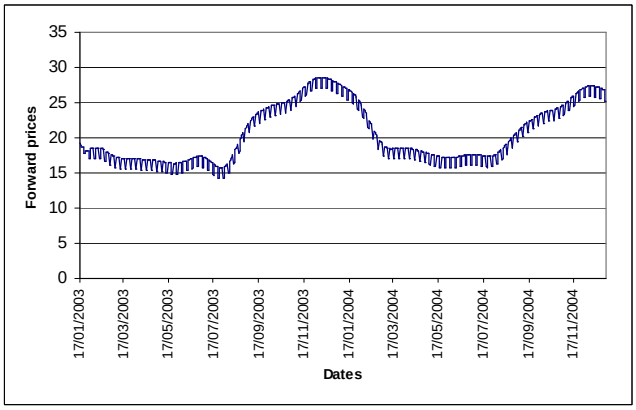

Figure 7: *Daily forward curve*

Let $P(N)$ be the price obtained for a quantization grid of size N, the error has been computed as $|P(N) - P(400)|, N \leq 400$. We assume that the error can be written as a functional of the grid size N with the following shape:

$$N \mapsto \frac{C}{N^{\alpha}}.$$

A linear regression in a logarithmic scale is done to find the functional that best fits the empirical error. The α coefficient obtained is 1.96.

Figures 8 and 9 show the obtained numerical convergence and the corresponding fitted functional.

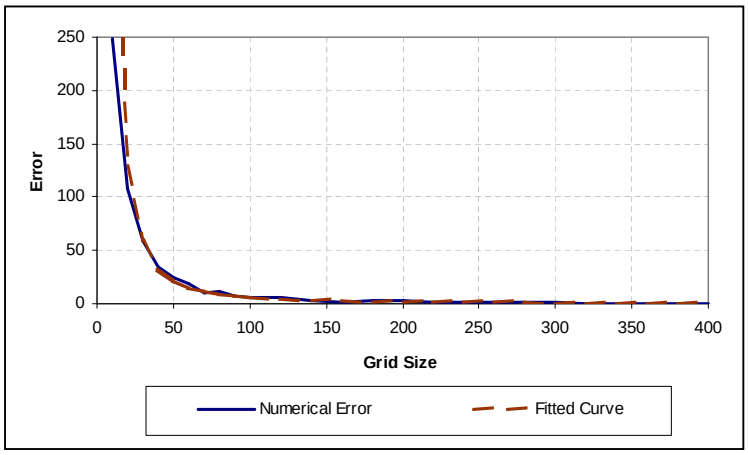

Figure 8: *Numerical Convergence*

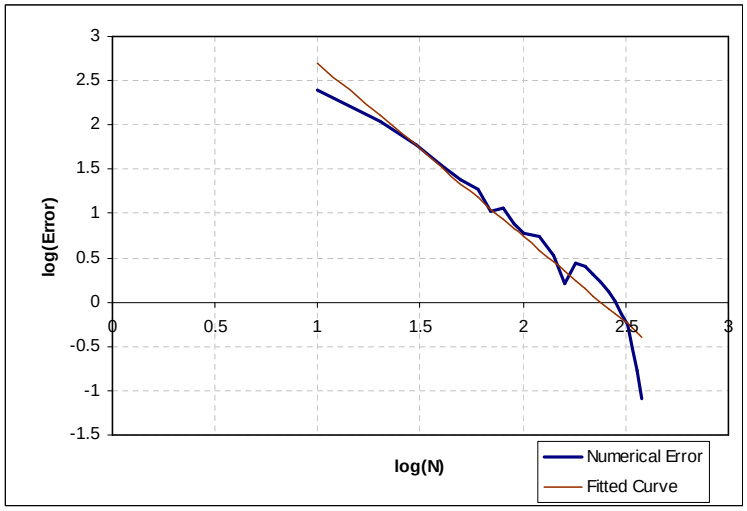

Figure 9: *Numerical Convergence (logarithmic scale)*

The same experiments have been done for other contracts, results are presented in Table 5. q_{min} and q_{max} are set to 0 and 6.

Forward Curve	Strike	Constraints $(Q_{min} - Q_{max})$	Estimated α
Figure 7	20	1300-1900	1.96
Flat (20)	20	1300-1900	2.07
Flat (20)	10	1300-1900	2.32
Flat (20)	20	1000-2000	1.95
Flat (20)	20	1600-1800	2.26

Table 5: *Estimation of the convergence rate*

We can conclude that the convergence rate of the quantization algorithm for pricing swing options is close to $O(\frac{1}{N^2})$. This convergence rate is much better than Monte-Carlo, and leads to

think that optimal quantization is an efficient alternative to Longstaff-Schwartz method for this problem.

4.2 Two factor model

We consider the following diffusion model for the forward contracts $(F_{t,T})_{0\leq t\leq T}$:

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma_1 e^{-\alpha_1(T-t)}dW_t^1 + \sigma_2 e^{-\alpha_2(T-t)}dW_t^2$$

where W^1 and W^2 are two Brownian motions with correlation coefficient ρ.

Standard computations based on Itô formula yield

$$S_t = F_{0,t}\exp\left(\sigma_1\int_0^t e^{-\alpha_1(t-s)}dW_s^1 + \sigma_2\int_0^t e^{-\alpha_2(t-s)}dW_s^2 - \frac{1}{2}\Lambda_t^2\right)$$

where

$$\Lambda_t^2 = \frac{\sigma_1^2}{2\alpha_1}(1-e^{-2\alpha_1 t}) + \frac{\sigma_2^2}{2\alpha_2}(1-e^{-2\alpha_2 t}) + 2\rho\frac{\sigma_1\sigma_2}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)t}).$$

Unlike the one factor model, the spot price process obtained from the two factor model is not a Markov process. Hence the dynamic programming equation (6) cannot be used directly. However, the structure process of the two factor model (See Annex, page 25)

$$X_t = \left(\int_0^t e^{-\alpha_1(t-s)}dW_s^1, \int_0^t e^{-\alpha_2(t-s)}dW_s^2\right)$$

is a Markov process, and $S_t = f(X_t)$ where $f : \mathbb{R}^2 \mapsto \mathbb{R}$ is a continuous function. So we can rewrite the dynamic programming equation as follows:

$$\begin{cases} P(t_k, X_{t_k}, Q_{t_k}) = \max_{q\in[q_{min},q_{max}]}\{q(f(X_{t_k})-K) + \mathbb{E}(P(t_{k+1}, X_{t_{k+1}}, Q_{t_k}+q)|X_{t_k})\}, \\ P(T, X_T, Q_T) = P_T(f(X_T), Q_T). \end{cases} \tag{13}$$

Then we need to quantize the $\mathbb{R}^2$ valued structure process $(X_k)_{k\geq 0}$. This can be done using the Fast Parallel Quantization (See Annex). The transitions are computed using Monte Carlo simulations and importance sampling (see (19)).

4.2.1 Call strip

We first consider a case without constraints, and compare the results with the theoretical price of the call strip, for several values of the strike K. The maturity of the contract is one month. Transitions have been computed with 3000000 of Monte-Carlo simulation. The parameters of the two factor model are:

$$\sigma_1 = 36\%, \alpha_1 = 0.21, \sigma_2 = 111\%, \alpha_2 = 5.4, \rho = -0.11. \tag{14}$$

Table 6 presents the results obtained for a strip of call. Even if the quantized process is taking values in $\mathbb{R}^2$, prices are close to the theoretical price even for small grids.

	$K = 5$	$K = 10$	$K = 15$	$K = 20$
Theoretical Price	2700	1800.21	924.46	268.59
50 point grid	2695.26	1795.26	918.18	261.06
100 point grid	2699.79	1799.89	923.64	267.17
200 point grid	2697.67	1797.83	921.94	266.71
300 point grid	2702.02	1800.16	924.40	268.52

Table 6: *Call Strip*

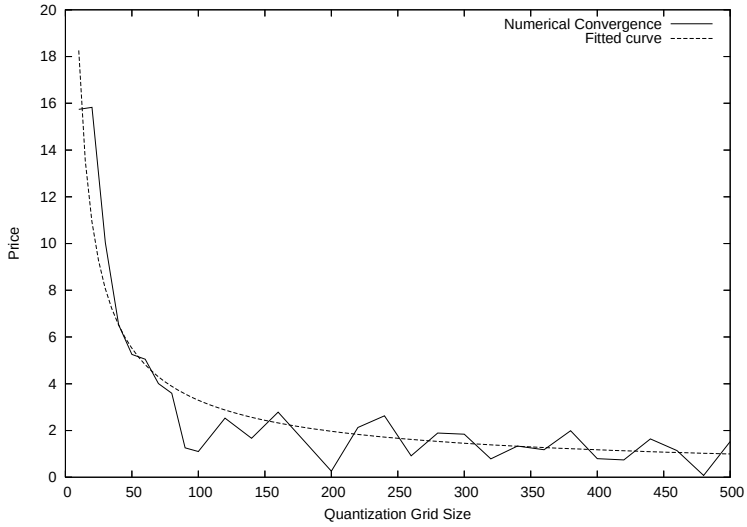

Figure 10: *Numerical Convergence*

Forward Curve	Strike	Constraints $(Q_{min} - Q_{max})$	Estimated α
Flat (20)	10	80-140	1.26
Flat (20)	20	80-140	1.00
Flat (20)	20	30-170	0.67
Flat (20)	20	100-120	1.19

Table 7: *Estimation of the convergence rate*

4.2.2 Convergence

We use the same procedure as in section 4.1.5 to find the functional $N \mapsto \frac{C}{N^\alpha}$ that best fits the empirical error.

Figure 10 shows an example of the empirical error and table 7 gather the values of α obtained for different contracts. The contract maturity has been set to one month.

The convergence of the quantization algorithm is close to $O(\frac{1}{N})$. The convergence rate is linked to the dimension d of the structure process, and from the results obtained in section 4.1.5 and in this section, we can assume that the convergence rate is close to $O(\frac{1}{N^{2/d}})$, which is better than the error bound theoretically established in Section 3.3.

4.3 Dimension reduction

In the case of multi-factorial models, we need to quantize the structure process $(X_k)_k$ instead of the spot process $S_{t_k} = f(X_k)$ in order to work with a Markov process (See section 4.2). Only the one factor model is Markovian. That is why quantization and Longstaff-Schwartz method have been compared just for this model. Longstaff Schwartz method also requires a Markov underlying process.

From an operational point of view, it is interesting to study the results obtained by formally quantizing the spot process (S_{t_k}), regardless to its dynamics, and using the approximation

$$\mathbb{E}(X|\mathcal{F}_{t_k}) \simeq \mathbb{E}(X|S_{t_k})$$

for any random variable X, even if the spot price (S_{t_k}) is not a Markov process. Similar approximation has already been proposed by Barraquand-Martineau in [5]. Numerical tests have shown that the resulting prices remain very close to those obtained by quantizing the structure process in the case of a two factor model. Execution time and convergence rate are significantly faster, and the quantization tree can be computed as presented in the Annex page 22 using Equation (21).

Even if there is no theoretical evidence on the error, this approach seems useful to get quick results. Table 8 presents some results, parameters of the two factor model are those of (14) and volume constraints are represented on Figure 4.

Quantized Process	$K = 5$	$K = 10$	$K = 15$	$K = 20$
Spot Price	30823.38	21414.30	13021.87	5563.68
Structure Process	30705.78	21518.75	13123.56	5722.89

Table 8: *Quantization of the spot price vs Quantization of the bivariate structure process*

The convergence rate obtained in this case is always $O(\frac{1}{N})$, which is consistent with the general rate $O(\frac{1}{N^{2/d}})$, because the quantized process is $\mathbb{R}$-valued.

Therefore, even if the spot process is not Markov, the quantization method can be performed all the same way as if it were, with small damage in practice. This is of course not a general consideration but rather an observation over the considered problem. The dramatic increase in the computation effort that can be gained from this observation can justify in this case a lack of formal rigor.

Conclusion

In this article, we have introduced an optimal quantization method for the valuation of swing options. These options are of great interest in numerous modeling issues of the energy markets and their accurate pricing is a real challenge.

Our method has been compared to the famous Longstaff-Schwartz algorithm and seems to perform much better on various examples. In fact, the optimal quantization method shares the good properties of the so called tree method but is not limited by the dimension of the underlying. Moreover, specific theoretical results provide *a priori* estimates on the error bound of the method.

Thus, optimal quantization methods suit very well to the valuation of complex derivatives and further studies should be done in order to extend the present results to other structured products arising in the energy sector.

References

[1] Bally V., Pagès G. (2001). A quantization algorithm for solving multi-dimensional optimal stopping problems, *Bernoulli*, **6**(9), 1-47.

[2] Bally V., Pagès G. (2003). Error analysis of the quantization algorithm for obstacle problems, Stochastic Processes & Their Applications, **106**(1), 1-40.

[3] Bally V., Pagès G., Printems J. (2005). A quantization tree method for pricing and hedging multi-dimensional American options, *Mathematical Finance*, **15**(1), 119-168.

[4] Bardou O., Bouthemy S., Pagès G. (2007). Pricing swing options using Optimal Quantization, pre-print LPMA-1141.

[5] Barraquand J. and Martineau D., (1995). Numerical valuation of high dimensional multivariate American securities, *Journal of Financial and Quantitative Analysis* **30**3, 383-405.

[6] Barrera-Esteve C., Bergeret F., Dossal C., Gobet E., Meziou A., Munos R., Reboul-Salze D. (2006). Numerical methods for the pricing of swing options: a stochastic control approach, *Methodology and Computing in Applied Probability.*

[7] Carmona R. and Ludkovski M. (2007). Optimal Switching with Applications to Energy Tolling Agreements, pre-print.

[8] Carmona R. and Ludkovski M. (2005). Gas Storage and Supply Guarantees : An Optimal Switching Approach, pre-print.

[9] Carmona R. and Dayanik S. (2003). Optimal Multiple-Stopping of Linear Diffusions and Swing Options, pre-print.

[10] Carmona R. and Touzi N. (2004). Optimal multiple stopping and valuation of swing options, *Mathematical Finance*, to appear.

[11] Clewlow L., Strickland C. and Kaminski V. (2002). Risk Analysis of Swing Contracts, *Energy and power risk management.*

[12] Clearwater S. H. and Huberman B. A. (2005). Swing Options : a Mechanism for Pricing It Peak Demand

[13] Eydeland A. and Wolyniec K. (2003). Energy and Power Risk Management, *Wiley Finance*

[14] Geman H. (2005). *Commodities and commodity derivatives - Modeling and Pricing for Agriculturals, Metals and Energy*, Wiley Finance.

[15] Graf S., Luschgy H. (2000). *Foundations of Quantization for Probability Distributions.* Lect. Notes in Math. 1730, Springer, Berlin.

[16] Jaillet P., Ronn E.I., Tompaidis (2004). Valuation of Commodity-Based Swing Options, *Management Science*, **50**, 909-921.

[17] Keppo J. (2004). Pricing of Electricity Swing Options, *Journal of Derivatives*, **11**, 26-43.

[18] Lari-Lavassani A., Simchi M., Ware A. (2001). A discrete valuation of Swing options, *Canadian Applied Mathematics Quarterly*, **9**(1), 35-74.

[19] F. Longstaff and E.S. Schwartz (2001). Valuing American Options by Simulation: A Simple Least Squares Approach, *The Review of Financial Studies*, **14**, 113-147.

[20] Meinshausen N. and Hambly B.M. (2004). Monte Carlo methods for the valuation of multiple exercise options, *Mathematical Finance* , **14**, $n^0 4$.

[21] Pagès G. (1998). A space vector quantization method for numerical integration, *J. Computational and Applied Mathematics*, 89, 1-38.

[22] Pagès G., Pham H., Printems J. (2004). Optimal quantization methods and applications to numerical problems in finance, *Handbook of Numerical Methods in Finance*, ed. S. Rachev, Birkhauser, 253-298.

[23] Pagès G. and Printems J. (2005). `www.quantize.maths-fi.fr`, website devoted to quantization, *maths-fi.com*

[24] Ross S.M. and Zhu Z. (2006). Structure of swing contract value and optimal strategy, *The 2006 Stochastic Modeling Symposium and Investment Seminar, Canadian Institute of Actuaries*

[25] Thompson A.C. (1995). Valuation of path-dependent contingent claims with multiple exercise decisions over time: the case of Take or Pay. *Journal of Financial and Quantitative Analysis*, **30**, 271-293.

[26] Winter C. and Wilhelm M. (2006). Finite Element Valuation of Swing Options, pre-print.

[27] Zador P.L. (1982). Asymptotic quantization error of continuous signals and the quantization dimension, *IEEE Trans. Inform. Theory*, **28**, Special issue on quantization, A. Gersho & R.M. Grey eds, 139-149.

Annex: Fast Parallel quantization (FPQ)

In this annex we propose an efficient method to quantize a wide family of spot price dynamics (S_{t_k}). To be precise we will assume that a time discretization Δ being fixed,

$$S_{k\Delta} = f_k(X_k), k \geq 0 \tag{15}$$

where $(X_k)_{k\geq 0}$ is a $\mathbb{R}^m$-valued Gaussian auto-regressive process and $(f_k)_{0\leq k\leq n}$ a family of continuous functions. The fast quantization method applies to the Gaussian process $(X_k)_{0\leq k\leq n}$. We will apply it to a scalar two factor model in full details. As a conclusion to this section we will sketch the approach to a multi-factor model.

Quantization of the Gaussian structure process

We consider a centered Gaussian first order auto-regressive process in $\mathbb{R}^m$:

$$X_{k+1} = AX_k + T\varepsilon_{k+1} \tag{16}$$

where $A \in \mathcal{M}(m \times m, \mathbb{R})$, $T \in \mathcal{M}(m \times m, \mathbb{R})$ lower triangular, and (ε_k) i.i.d. with $\mathcal{N}(0, I_m)$ distribution.

Denote by $D(Z) = [\mathbb{E}(Z_i Z_j)]_{1\leq i,j\leq m}$ the covariance matrix of Z. We have $\forall k \in \mathbb{N}$:

$$D(X_{k+1}) = AD(X_k)A^* + TT^*.$$

Denote Σ_k the lower triangular matrix such that $D(X_k) = \Sigma_k \Sigma_k^*$.

We consider for every $k = 0, \ldots, n-1$, an *optimal (quadratic) quantizer* $x_k^{(N_k)}$ of size N_k, for the $\mathcal{N}(0, I_m)$ distribution. The quantization grid of the random variable X_k is taken as a dilatation of $x_k^{(N_k)}$,*i.e*

$$\bar{x}_k = \Sigma_k \, x_k^{(N_k)} := (\Sigma_k x^{(N_k),i})_{1 \le i \le N_k}.$$

To calculate the conditional expectations in the dynamic programming equation, we need to get the following transition probabilities:

$$\pi_k^{ij} = \mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}) | X_k \in C_i(\bar{x}_k))$$

where $C_i(x)$ denotes the i-th Voronoi cell of the generic quantizer $x \in (\mathbb{R}^d)^N$. Then

$$\mathbb{P}(X_k \in C_i(\bar{x}_k)) = \mathbb{P}(Z \in C_i(x_k^{(N_k)})),$$

with $Z \sim \mathcal{N}(0, I_m)$. This probability is provided as a companion parameter with the normal distribution grid files (available on [23]).

To get the transition probability π_k^{ij} we need to compute

$$p_k^{ij} = \mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}), X_k \in C_i(\bar{x}_k)).$$

Proposition 4.1. *Let X be a discrete time process described as above. Let U, V be two gaussian random variables $\mathcal{N}(0, I_m)$. Then we have for every $k \in \{0, \ldots, n-1\}$, every $i \in \{1, \ldots, N_k\}$, every $j \in \{1, \ldots, N_{k+1}\}$,*

$$\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}), X_k \in C_i(\bar{x}_k)) = \mathbb{P}(U \in C_i(x_k^{(N_k)}), A_{k+1}U + B_{k+1}V \in C_j(x_{k+1}^{N_{k+1}})) \qquad (17)$$

where A_k and B_k are $q \times q$ matrices whose coefficients depend on k, and on the matrices A, T. If $k = 0$,

$$\mathbb{P}(X_1 \in C_j(\bar{x}_1)) = \mathbb{P}(V \in C_j(x_1^{(N_1)})). \qquad (18)$$

Proof. We have:

$$\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)) = \mathbb{P}(AX_k + T\varepsilon_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)).$$

We consider the couple $(X_k, T\varepsilon_{k+1})$. $T\varepsilon_{k+1}$ is independent of X_k. Let $\eta = (\eta_1, \eta_2)$ a couple of independent Gaussian random vectors: $\eta_i \sim N(0, I_m), i = 1, 2$. Then $(X_k, T\varepsilon_{k+1}) \sim (\Sigma_k \eta_1, T\eta_2)$ and

$$\begin{aligned}
\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)) &= \mathbb{P}(A\Sigma_k \eta_1 + T\eta_2 \in C_j(\bar{x}_{k+1}); \Sigma_k \eta_1 \in C_i(\bar{x}_k)) \\
&= \mathbb{P}(\Sigma_{k+1}^{-1}(A\Sigma_k \eta_1 + T\eta_2) \in C_j(x_{k+1}^{(N_{k+1})}); \eta_1 \in C_i(x_k^{(N_k)})).
\end{aligned}$$

Setting

$$A_{k+1} = \Sigma_{k+1}^{-1} A \Sigma_k, \quad B_{k+1} = \Sigma_{k+1}^{-1} T$$

we get

$$\mathbb{P}(X_{k+1} \in C_j(\bar{x}_{k+1}); X_k \in C_i(\bar{x}_k)) = \mathbb{P}(A_{k+1}\eta_1 + B_{k+1}\eta_2 \in C_j(x_{k+1}^{(N_{k+1})}); \eta_1 \in C_i(x_k^{(N_k)})).$$

If $k = 0$ and $\Sigma_0 \equiv 0$, the quantity

$$\mathbb{P}(X_k \in C_j(\bar{x}_k)) = \mathbb{P}(\eta_1 \in C_j(x_k^{(N_k)}))$$

is given as a companion parameter with the quantization grids of the normal distribution. □

Remark 4.1. Equation (17) emphasizes the fact that the transitions can be computed in parallel.

Remark 4.2. To simplify the structure of the quantization tree we propose to consider the same normalized grid of size $N_k = N$ at each step k but other choices are possible like those recommended in [1].

Numerical methods

Hereafter we will focus on the numerical computation of these transitions.

• The standard Monte Carlo approach The simplest way is to use a Monte Carlo method. One just needs to simulate couples of independent gaussian random variables (η_1, η_2). This approach can be used whatever the dimension m of the random variables η_1 and η_2 is. It can clearly be parallelized as any MC simulation but fail to estimate the transition form states which are not often visited.

• Fast Parallel Quantization Method In order to improve the accuracy, especially for the points $x^{(N_{k+1}),j}$ of the grids $x_{k+1}^{(N_{k+1})}$ which are rarely reached by the paths starting from the cell of $x^{(N_k),i}$, it is possible to perform importance sampling. The idea is to use Cameron-Martin formula to re-center the simulation: for every $k \in \{0, \dots, n-1\}$ and every $i \in \{1, \dots, N_k\}$,

$$p_k^{ij} = \mathbb{P}\left(\eta_1 \in C_i(x_k^{(N_k)}); A_{k+1}\eta_1 + B_{k+1}\tilde{\eta}_2 \in C_j(x_{k+1}^{(N_{k+1})})\right) =$$
$$e^{-\frac{1}{2}|x_k^{(N_k),i}|^2}\mathbb{E}\left(e^{-(x_k^{(N_k),i})^*\eta_1}\mathbf{1}_{\{\eta_1 + x_k^{(N_k),i} \in C_i(x_k^{(N_k)})\}}\mathbf{1}_{\{A_{k+1}(\eta_1 + x_k^{(N_k),i}) + B_{k+1}\eta_2 \in C_j(x_{k+1}^{(N_{k+1})})\}}\right). \quad (19)$$

Then these expectations can be computed by Monte Carlo simulations, the transitions between the different times steps can be computed in parallel.

• Quantized Parallel Quantization Method If $m = 1$, the transitions can be computed using again optimal quantization, because in low dimension (say $d \leq 4$), quantization converges faster than Monte Carlo method. In this case, we have to compute a two dimensional expectation.

We estimate for every $k \in \{0, \dots, n-1\}$, every $i \in \{1, \dots, N_k\}$ and every $j \in \{1, \dots, N_{k+1}\}$ the following probabilities:

$$p_k^{ij} = \mathbb{P}(\eta_1 \in C_i(x_k^{(N_k)}); \alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(x_{k+1}^{(N_{k+1})})) \quad (20)$$

where $(\eta_1, \eta_2) \sim \mathcal{N}(0, I_2)$, and α_k and β_k are scalar coefficients satisfying $\alpha_k^2 + \beta_k^2 = 1$:

$$\begin{aligned} \alpha_k &= \tfrac{\Sigma_k A}{\Sigma_{k+1}} \\ \beta_k &= \tfrac{T}{\Sigma_{k+1}}. \end{aligned}$$

To alleviate notations, we temporarily set $x = x_k^{(N_k)}$ and $y = x_{k+1}^{(N_{k+1})}$.

We define $\left[x^{i-\frac{1}{2}}, x^{i+\frac{1}{2}}\right] = \left[\frac{1}{2}(x^i + x^{i-1}), \frac{1}{2}(x^i + x^{i+1})\right] = C_i(x)$ (the same shortcut is implicitly defined for y).

In order to reduce the problem dimension, it is possible to write the probability p_{ij}^k as a double integral, and to integrate first with respect to the second variable by using Fubini theorem:

$$\mathbb{P}(\eta_1 \in C_i(x);\ \alpha_{k+1}\eta_1 + \beta_{k+1}\eta_2 \in C_j(y)) =$$
$$\mathbb{E}\left(\mathbf{1}_{\{x^{i-\frac{1}{2}} \leq \eta_1 \leq x^{i+\frac{1}{2}}\}}\left(\mathcal{N}(\frac{y^{j-\frac{1}{2}} - \alpha_{k+1}\eta_1}{\beta_{k+1}}) - \mathcal{N}(\frac{y^{j+\frac{1}{2}} - \alpha_{k+1}\eta_1}{\beta_{k+1}})\right)\right)$$

where $\mathcal{N}(x)$ is the distribution function of the normal distribution.

Importance sampling can again be used to improve the results precision. Eventually we have to compute the following one dimensional expectation:

$$\mathbb{P}\left(\eta_1 \in C_i(x); \alpha_{k+1}\eta_1 + \beta_{k+1}\tilde{\eta}_2 \in C_j(y)\right) = e^{\frac{-(x^i)^2}{2}}\mathbb{E}\left[e^{-x^i\eta_1}\mathbf{1}_{\left\{-\frac{\Delta x^i}{2}\leq\eta_1\leq\frac{\Delta x^{i+1}}{2}\right\}}\left\{\mathcal{N}\left(\frac{y^{j-\frac{1}{2}} - \alpha_{k+1}(\eta_1 + x^i)}{\beta_{k+1}}\right) - \mathcal{N}\left(\frac{y^{j+\frac{1}{2}} - \alpha_{k+1}(\eta_1 + x^i)}{\beta_{k+1}}\right)\right\}\right]. \quad (21)$$

For this one-dimensional expectation computation, quantization can be used again since it converges faster than Monte Carlo method.

Example: Two factor model

We consider the following diffusion model for the forward contracts $(F_{t,T})_{0\leq t\leq T}$:

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma_1 e^{-\alpha_1(T-t)}dW_t^1 + \sigma_2 e^{-\alpha_2(T-t)}dW_t^2$$

where W^1 and W^2 are two Brownian motions with correlation coefficient ρ.

Standard computations based on Itô formula yield

$$S_t = F_{0,t}\exp\left(\sigma_1\int_0^t e^{-\alpha_1(t-s)}dW_s^1 + \sigma_2\int_0^t e^{-\alpha_2(t-s)}dW_s^2 - \frac{1}{2}\Lambda_t^2\right)$$

where

$$\Lambda_t^2 = \frac{\sigma_1^2}{2\alpha_1}(1-e^{-2\alpha_1 t}) + \frac{\sigma_2^2}{2\alpha_2}(1-e^{-2\alpha_2 t}) + 2\rho\frac{\sigma_1\sigma_2}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)t}).$$

We have $S_t = F_{0,t}\exp\left(\sigma_1 X_t^1 + \sigma_2 X_t^2 - \frac{1}{2}\Lambda_t^2\right)$, where X_t is the following structure process:

$$X_t = \left(\int_0^t e^{-\alpha_1(t-s)}dW_s^1, \int_0^t e^{-\alpha_2(t-s)}dW_s^2\right). \quad (22)$$

Proposition 4.2. *Let $Z = (Z_t)$ be an Ornstein-Uhlenbeck process. $Z_t = Z_0 + \int_0^t e^{-\alpha(t-s)}dB_s$ where B is a standard Brownian motion, and Z_0 is Gaussian and independent of B. Z can be written at discrete times $k\Delta$ as a first order auto-regressive process:*

$$Z_{k+1} = e^{-\alpha\Delta}Y_k + \sqrt{1-e^{-2\alpha\Delta}}\sqrt{\frac{1}{2\alpha}}\varepsilon_{k+1} \quad (23)$$

where (ε_k) is i.i.d and $\varepsilon_1 \sim \mathcal{N}(0,1)$.

X_t is made up with two Ornstein-Uhlenbeck processes. Using Proposition 4.2, it yields:

Proposition 4.3.

$$X_{k+1} = AX_k + T\varepsilon_{k+1}$$

with $\varepsilon_k \sim \mathcal{N}(0, I_2)$ i.i.d. and

$$A = \begin{bmatrix} e^{-\alpha_1\Delta} & 0 \\ 0 & e^{-\alpha_2\Delta} \end{bmatrix}$$

$$r = \rho \frac{\frac{1}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta})}{\sqrt{\frac{1}{4\alpha_1\alpha_2}(1-e^{-2\alpha_1\Delta})(1-e^{-2\alpha_2\Delta})}}$$

$$T = \begin{bmatrix} \frac{1}{2\alpha_1}(1-e^{-2\alpha_1\Delta}) & 0 \\ \frac{1}{2\alpha_2}(1-e^{-2\alpha_2\Delta})r & \frac{1}{2\alpha_2}(1-e^{-2\alpha_2\Delta})\sqrt{1-r^2} \end{bmatrix}.$$

Hence it is possible to use the fast parallel quantization method described in section 4.3.

General multi-factor Gaussian model

More generally, we consider a family of price dynamics that can be written as follows:

$$\frac{dF_{t,T}}{F_{t,T}} = \sum_{i=1}^{m} P_i(T-t)e^{-\alpha_i(T-t)}dW_t^i \tag{24}$$

where $P_i(x)$ is a polynomial function of degree d_i, for every $i = 1, \ldots, m$, and W is a Brownian motion, with $d < W^i, W^j >_t = \rho_{ij}dt$.

The two factor model (Section 4.3) corresponds to $m = 2$, $P_i \equiv \sigma_i$, $i = 1, 2$.

In order to price a swing option with such a model, we first need to quantize it. Equation (24) yields:

$$F_{t,T} = F_{0,T}e^{\sum_{i=1}^{m}\int_0^t P_i(T-s)e^{-\alpha_i(T-s)}dW_s^i - \frac{1}{2}\phi(t,T)} \tag{25}$$

where

$$\phi(t,T) = \sum_{i=1}^{m}\int_0^t P_i^2(T-s)e^{-2\alpha_i(T-s)}ds + \sum_{i\neq j}\rho_{ij}\int_0^t P_i(T-s)P_j(T-s)e^{-(\alpha_i+\alpha_j)(T-s)}ds.$$

Practically we focus on the spot price $F_{t,t}$ or the day-ahead contract $F_{t,t+1}$. Unfortunately these processes are not Markovian in a general setting, except when $m = 1$ and $d_1 = 0$ (Ornstein-Uhlenbeck process).

We consider a discretization time step $\Delta > 0$, and we set, for all $i \in \{1, \ldots, m\}$ and for all $l \in \{0, \ldots, d_i\}$

$$X_k^{i,l} = \int_0^{k\Delta} P_i\left((k+l)\Delta - s\right)e^{-\alpha_i((k+l)-s)}dW_s^i.$$

Proposition 4.4. *$X_k = [X_k^{i,l}]_{1\leq i\leq m, 0\leq l\leq d_i}$ is a $\mathbb{R}^{d_1+\cdots+d_m+m}$-valued gaussian AR(1).*

Lemma 4.5. *Let $P \in \mathbb{R}[Z]$, $d^oP = d$ and $\theta \in \mathbb{R}^*$. Then $(P(Z+l\theta))_{0\leq l\leq d}$ is a basis of $\mathbb{R}^d$.*

Proof. Using a dimension argument, only the linear independence of the family has to be checked. And we have

$$\begin{aligned} \sum_{k=0}^{d}\lambda_k P(Z+k\theta) = 0 \quad &\Leftrightarrow \quad \sum_{k=0}^{d}\lambda_k\sum_{j=0}^{d}\frac{(k\theta)^j}{j!}P^{(j)}(Z) = 0 \\ &\Leftrightarrow \quad \sum_{j=0}^{d}\frac{\theta^j}{j!}\left(\sum_{k=0}^{d}\lambda_k k^j\right)P^{(j)}(Z) = 0. \end{aligned}$$

Since $\left(P^{(j)}(Z)\right)_{0\leq j\leq d}$ is a basis of $\mathbb{R}_d[Z]$, it yields

$$\forall j \in \{0, \ldots, d\}, \quad \sum_{k=0}^{d}\lambda_k k^j = 0$$

so that $\lambda_k = 0,\ 0 \leq k \leq d$ since $\det[k^j]_{0\leq k,j\leq d} \neq 0$ (Vandermonde determinant). □

Proof. (of Proposition 4.4) We can extend the definition of $X_k^{i,l}$ to $l \in \mathbb{N}$. It is easy to check that

$$\begin{aligned} X_{k+1}^{i,l} &= \int_0^{(k+1)\Delta} P_i\left((k+1+l)\Delta - s\right) e^{-\alpha_i((k+1+l)\Delta - s)} dW_s^i \\ &= X_k^{i,l+1} + \varepsilon_{k+1}^{i,l} \end{aligned}$$

where $\varepsilon_{k+1}^{i,l} = \int_{k\Delta}^{(k+1)\Delta} P_i\left((k+1+l)\Delta - s\right) e^{-\alpha_i((k+1+l)\Delta - s)} dW_s^i$.

If $l = d_i$,

$$X_{k+1}^{i,d_i} = X_k^{i,d_i+1} + \varepsilon_{k+1}^{i,d_i}.$$

According to Lemma 4.5,

$$P_i\left(Z + (d_i+1)\Delta\right) = \sum_{l=0}^{d_i} \lambda^{i,l} P_i(Z + l\Delta).$$

Hence

$$\begin{aligned} X_k^{i,d_i+1} &= \int_0^{k\Delta} P_i\left((k+d_i+1)\Delta - s\right) e^{-\alpha_i((k+d_i+1)\Delta - s)} dW_s^i \\ &= \sum_{l=0}^{d_i} \lambda^{i,l} \int_0^{k\Delta} P_i\left((k+l)\Delta - s\right) e^{-\alpha_i((k+l)\Delta - s)} dW_s^i e^{-\alpha_i(d_i+1-l)\Delta} \\ &= \sum_{l=0}^{d_i} \lambda^{i,l} e^{-\alpha_i(d_i+1-l)\Delta} X_k^{i,l} \\ &= \sum_{l=0}^{d_i} \tilde{\lambda}^{i,l} X_k^{i,l}. \end{aligned}$$

Finally we have

$$X_{k+1}^{i\cdot} = A^i X_k^{i\cdot} + \varepsilon_{k+1}^{i\cdot}$$

where

$$A^{i\cdot} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ \tilde{\lambda}^{i,0} & \cdots & \cdots & \cdots & \tilde{\lambda}^{i,d_i} \end{pmatrix}$$

and

$$X_{k+1} = AX_k + \varepsilon_k, X_0 = 0$$

where $\varepsilon_k \in \sigma\left(W_u^i - W_{k\Delta}^i, k\Delta \le u \le (k+1)\Delta, i = 1, \ldots, m\right)$ is independent of $\mathcal{F}_{k\Delta}^W$. The process $(X_k)_k$ is thus a gaussian AR(1). □

X_k is the structure process for the spot price $S_{k\Delta} = F_{k\Delta,k\Delta}$. Its dimension is $\sum_{i=1}^m (d_i+1)$. For the two factor model, the structure process is $\mathbb{R}^2$-valued, because $m = 2$, and $d_i = 0, i = 1, 2$. This is coherent with (22).